MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

and

CENTER FOR BIOLOGICAL INFORMATION PROCESSING
WHITAKER COLLEGE

A.I. Memo 777
C.B.I.P. Paper 005

May, 1984

A GENERALIZED ORDERING CONSTRAINT FOR STEREO CORRESPONDENCE

A. L. Yuille and T. Poggio

Abstract. The ordering constraint along epipolar lines is a powerful constraint that has been exploited by some recent stereomatching algorithms. We formulate a *generalized ordering constraint*, not restricted to epipolar lines. We prove several properties of the generalized ordering constraint and of the "forbidden zone", the set of matches that would violate the constraint. We consider both the orthographic and the perspective projection case, the latter for a simplified but standard stereo geometry. The disparity gradient limit found in the human stereo system may be related to a form of the ordering constraint. To illustrate our analysis we outline a simple algorithm that exploits the generalized ordering constraint for matching contours of wireframe objects. We also show that the use of the generalized ordering constraint implies several other stereo matching constraints: a) the ordering constraint along epipolar lines, b) figural continuity, c) Binford's cross-product constraint, d) Mayhew and Frisby's figural continuity constraint. We finally discuss ways of extending the algorithm to arbitrary 3-D objects.

This report describes research done within the Artificial Intelligence Laboratory and the Center for Biological Information Processing (Whitaker College) at the Massachusetts Institute of Technology. Support for the A. I. Laboratory's research in artificial intelligence is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-80-C-0505. The Center's support is provided in part by the Sloan Foundation and in part by Whitaker College.

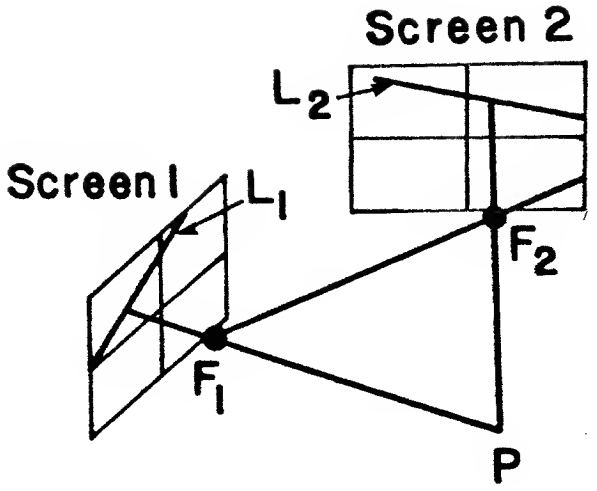

Figure 1 The geometry of the epipolar lines. *The plane defined by the two foci, F_1 and F_2, and a point P cuts the two image planes in the corresponding epipolar lines l_1 and l_2. As we vary the position of P we produce a 1-parameter family of epipolar lines on each screen. The lines are not parallel to one another, and it can be seen that the epipolar lines on screen 1 all intersect at the point where the line F_1F_2 hits screen 1.*

1. Introduction

The problem of stereo matching is ill-posed and underdetermined: constraints are needed (a) to make the solution unique and (b) to reduce the search problem among possible matches.

Marr and Poggio (1979) originally identified two important constraints: (1) *uniqueness*, that is, an element in one image in general only corresponds with a single element in the other image, and (2) *continuity*, that is, stereo disparity varies smoothly almost everywhere in the image. These constraints are powerful because they do not depend on the specific properties of the scene but on general properties of the stereo geometry. Marr and Poggio (1979) proposed a stereo matching algorithm, further developed by Grimson (1981, 1984), which incorporates the uniqueness and continuity constraints to match zero-crossing descriptions computed at different scales. An *ordering constraint along epipolar lines* has been exploited, both implicitly and explicitly, in several computer algorithms for stereo matching, as a special instance of the continuity constraint. Epipolar lines in the two images are lines on which corresponding points lie. The projections of a point P in space lie on the plane defined by P and the two camera foci and, as a consequence, on the two lines defined by the intersection of this plane with the two image planes (see figure 1).

This implies that the matching problem can be reduced to a one-dimensional search if the epipolars are known. Most algorithms assume that the epipolar geometry is known (from a known camera geometry) and that the images are registered. Furthermore, the *ordering* of edges or other features is usually preserved by stereo projection along epipolar lines (that is, if feature A is to the left of feature B in the left stereo image, then this spatial relationship is maintained in the right stereo image). The ordering constraint along epipolar lines follows from the continuity of surfaces and the assumption of opacity. As originally suggested by Baker (1982) the ordering constraint is violated in situations such as figure 2. Recently, Verri (1984) has discussed the role of the "forbidden zone", where the ordering constraint is violated (Krol and van de Grind, 1982 first introduced the notion of forbidden zone). The forbidden zone associated with each point of the visible surface is a set of points in space that would have images violating the ordering constraint. If any point in the forbidden zone

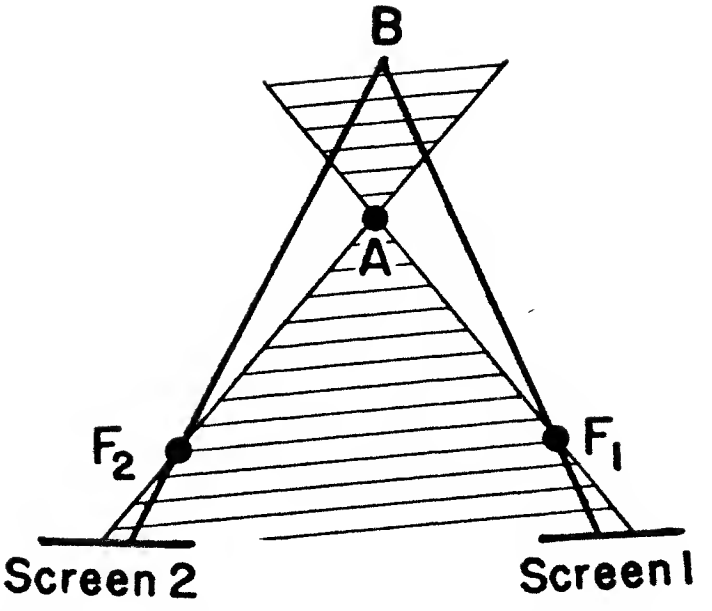

Figure 2 *The forbidden zone associated with point A is represented by the dashed region. Any point in this region such as B has projections which violate the ordering constraint relative to A.*

would be connected to the first point by an opaque surface the two images would "see" opposite sides of the surface.

This *ordering constraint* can be exploited to reduce the complexity of the search for matching features, and to eliminate false matches. Interestingly, there is preliminary evidence that the ordering contraint (and perhaps a stronger form of it) may be implemented in the human stereo system (Burt & Julesz 1980). The human system, however, must often cope with situations in which the images are not precisely registered. Furthermore, physical edges are inherently two-dimensional (a property that is not exploited by the epipolar ordering constraint). It is therefore natural to ask whether the ordering constraint can be generalized from epipolar lines[1]. More precisely, can an ordering constraint be formulated that is independent of the epipolar geometry? The simplest use of such a constraint would be when either the epipolar lines are unknown or their estimation is affected by errors.

In this paper we show that it is indeed possible to generalize the ordering constraint. We also give an analytic definition of the forbidden zone and characterize its properties. We show then that the generalized ordering constraint implies several other constraints that have been exploited in stereo matching. An algorithm based on this constraint has not been implemented yet, but we discuss its advantages and its limitations.

The plan of the paper is as follows. Throughout the paper, unless specifically stated, we consider a stereo geometry in which the two image planes have the same vertical unit vectors. This simple geometry is fully representative for most practical applications and it represents a good approximation for human binocular geometry (Longuet-Higgins, 1982). In section 2 we assume orthographic projection. We prove a simple relationship between the two images of a 3D curve that leads to a generalization of the standard ordering constraint. This relationship allows us to identify special points in the images that correspond uniquely to the same physical point in the object curve. The Generalized Ordering Constraint (GOC) implies several of the specific constraints listed by Baker et al. (1983) , Mayhew and Frisby (1981) (see their figural continuity constraint), and Ohta and Kanade (1983). The ordering constraint breaks down when the object curve enters the *forbidden zone*. We define the

[1] This question was first brought to our attention by Dr. V. Torre

forbidden zone geometrically and algebraically and discuss its properties. Section 3 shows that slightly weaker results can be derived for perspective projection (this includes the ordering constraint along epipolar lines as a very special case). These results again break down for the forbidden zone, which is more complex than in the orthographic case. In section 3, we characterize fully the forbidden zone for the perspective case and prove several results about it. The boundaries of the forbidden zone correspond to the so called Panum limiting case in the psychophysics of stereo, when one line of sight just grazes the surface. Burt and Julesz' results suggest that human stereovision is limited to a smaller disparity gradient. In section 4 we derive the equations for the physical surfaces corresponding to this *disparity gradient* limit. In section 5 we show how the generalized ordering constraint implies other stereo matching constraints. In section 6 we outline an algorithm for stereo based on matching contours. From a single contour the algorithm retrieves the viewing parameters and unambiguously matches points along the contour using the generalized ordering constraint. The constraint of figural continuity (Mayhew and Frisby, 1981) follows from the generalized ordering constraint and is implicitly implemented in our algorithm. In Appendix 2 we derive an explicit solution of the equations for the viewing parameters in the case of the stereo geometry suggested by Longuet-Higgins, using only two points.

2. The Ordering Constraint and the Forbidden Zone for Orthographic Projection

We assume a stereo-imaging geometry of the type proposed by Longuet-Higgins (1982) in which the planes of the horizontal meridians of the two eyes coincide. Hence, the relative orientation of the two eyes is defined by one parameter only. We also assume orthographic projection. We show that for any part of the object which does not lie in the forbidden zone, there is a simple relationship between the images of the object in the two eyes. This enables us to generalize the ordering constraint (Baker & Binford, 1981) and to identify features in the images that correspond to the same feature in the object.

We define an orthonormal triad of vectors in each eye. By our restriction on the geometry, the vertical direction, **k**, is the same for each eye. The right eye has vectors **i**, **j** and **k**, where **j** is normal to the right image plane. Similarly the left eye has vectors **i′**, **j′** and **k**, where **j′** is normal to the left image plane The convergence angle ϑ satisfies $\mathbf{j} \cdot \mathbf{j}' = \cos\vartheta$ (see Figure 3). We define coordinates X, Y and Z along the **i**, **j** and **k** axes respectively. Similarly, we let X' and Y' be coordinates along the **i′** and **j′** directions. Note that the origins of these systems of coordinates lies at the intersection of the two image planes and not at the focal points of the eyes. The following equations connect the triads of the two eyes:

$$\begin{aligned} \mathbf{j} \cdot \mathbf{j}' &= \cos\vartheta \\ \mathbf{i} \cdot \mathbf{i}' &= \cos\vartheta \\ \mathbf{i}' \cdot \mathbf{j} &= -\sin\vartheta \\ \mathbf{i} \cdot \mathbf{j}' &= +\sin\vartheta \end{aligned} \tag{2.1}$$

We consider now a curve in the right image plane, parameterized by s (which is *not* the arc length) and written as

$$\mathbf{r}_R(s) = X(s)\mathbf{i} + Z(s)\mathbf{k} \tag{2.2}$$

Under the assumption of orthographic projection, equation (2.2) is the image of an object curve given by $\mathbf{r}(s) = X(s)\mathbf{i} + Y(s)\mathbf{j} + Z(s)\mathbf{k}$,

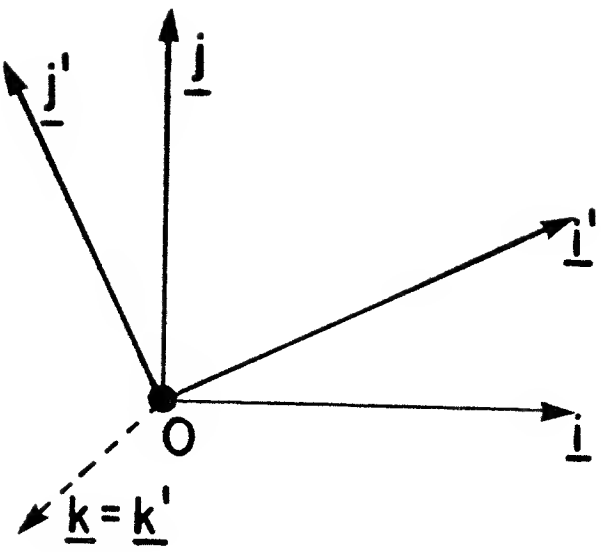

Figure 3 *See text.*

$$r_L(s) = X'(s)\mathbf{i}' + Z'(s)\mathbf{k} \tag{2.3}$$

where

$$\begin{aligned} X'(s) &= \mathbf{r}(s)\cdot\mathbf{i}' = X(s)\cos\vartheta - Y(s)\sin\vartheta \\ Z'(s) &= \mathbf{r}(s)\cdot\mathbf{k} = Z(s) \end{aligned} \tag{2.4}$$

We partition the right image curve into intervals for which Z is a single valued function of X. These intervals are separated by points at which $\frac{dX}{dZ} = 0$. For each such interval X can be chosen as the parameter and we can write the projection (in the right image plane) as

$$\mathbf{r}_R(X) = X\mathbf{i} + Z(X)\mathbf{k} \tag{2.5}$$

Identification of s in equation (2.3) with X yields (for the left image)

$$\mathbf{r} = X'(X)\mathbf{i}' + Z'(X)\mathbf{k} \tag{2.6}$$

with

$$\begin{aligned} X'(X) &= X\cos\vartheta - Y(X)\sin\vartheta \\ Z'(X) &= Z(X) \end{aligned} \tag{2.7}$$

Let us now compute $\frac{dX'}{dX}$ from equation (2.4):

$$\frac{dX'}{dX} = \cos\vartheta - \frac{dY(X)}{dX}\sin\vartheta \tag{2.8}$$

Note that (see Figure 4)

$$\frac{dY}{dX} < \frac{\cos\vartheta}{\sin\vartheta} \tag{2.9}$$

is the condition that curve never enters the "forbidden zone". Thus, if the curve never enters the forbidden zone, i.e., it satisfies the condition (2.9), *then*

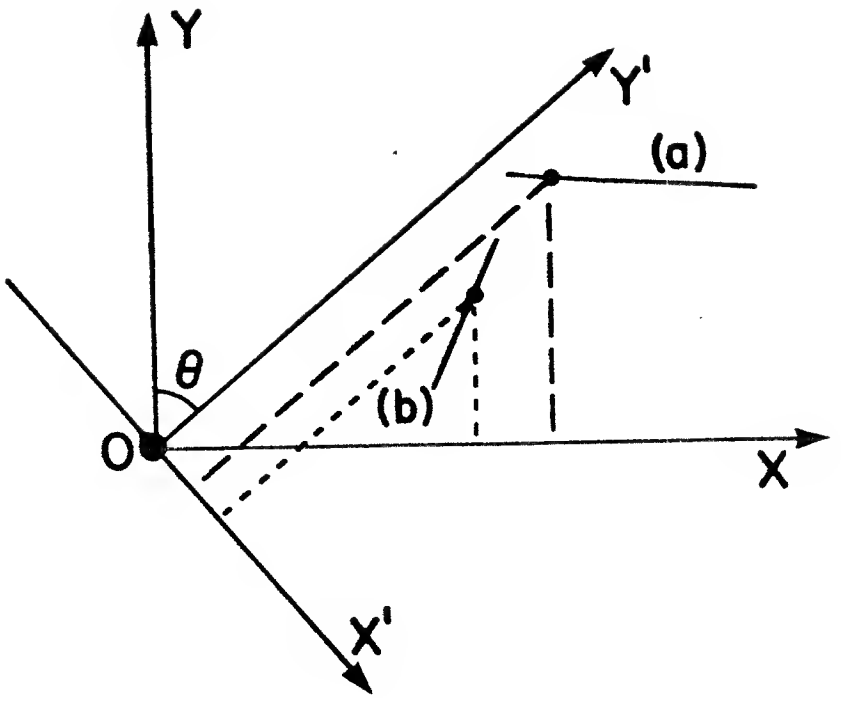

Figure 4 *Line (a) has gradient smaller than cotϑ and therefore is not in the forbidden zone. Line (b) has gradient larger than cotϑ and lies in the forbidden zone. The two views are of "opposite" sides.*

$$\frac{dX'}{dX} > 0. \tag{2.10}$$

Let us consider now the slope of the curve in the two images for each partition. The chain rule yields

$$\frac{dZ'}{dX'} = \frac{dZ}{dX}\frac{dX}{dX'} \tag{2.11}$$

Thus, if the curve is outside the forbidden zone, (a) $\frac{dZ'}{dX'}$ and $\frac{dZ}{dX}$ have the same sign (because of equation (2.10). Moreover, (b) the zeros of $\frac{dZ'}{dX'}$ correspond to zeros of $\frac{dZ}{dX}$. Furthermore, (c) equation (2.10) gives an ordering constraint on matching corresponding points in the two curves, since within each partition ordered points on the right image curve correspond to points on the right image curve with the same order (because of monotonicity implied by equation (2.11)).

The Forbidden Zone for Orthographic Projection

We now define the forbidden zone and show that it occurs precisely where the ordering constraints break down.

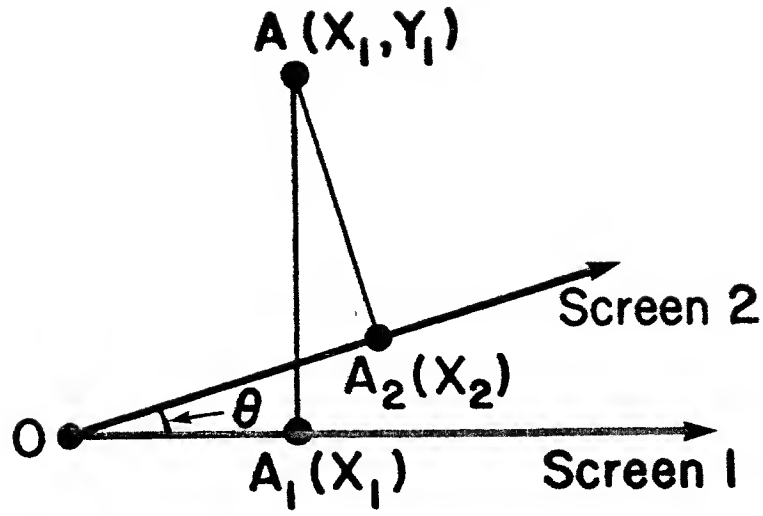

Figure 5 *Points A_1 and A_2 are the images of the physical point A and have coordinates X_1 and X_2, respectively.*

Suppose the two screens, screen 1 and screen 2, are at an angle θ to each other as illustrated in Figures 5 and 6. Consider two points A and A' with coordinates (X_1, Y_1) and (X_1', Y_1') respectively relative to screen 1. They will be projected to points A_1 and A_1' on screen 1 with $OA_1 = X_1$ and $OA_1' = X_1'$. They will be projected to points A_2 and A_2' on screen 2 with $OA_2 = X_2 = X_1 \cos\theta + Y_1 \sin\theta$ and $OA_2' = X_2' = X_1' \cos\theta + Y_1' \sin\theta$. The ordering constraint will hold, provided

$$X_1' - X_1 > 0, \text{if and only if} X_2' - X_2 > 0 \tag{2.12}$$

We have

$$X_2' - X_2 = (X_1' - X_1)\cos\theta\{1 + \frac{Y_1' - Y_1}{X_1' - X_1}\tan\theta\} \tag{2.13}$$

So a necessary and sufficient condition for (2.12) to hold is that

$$1 + \frac{Y_1' - Y_1}{X_1' - X_1}\tan\theta > 0 \tag{2.14}$$

Note that we have $-\pi/2 < \theta < \pi/2$ and hence $\cos\theta$ is always positive.

Equation (2.14) can be interpreted as a condition on the gradient m of the straight line joining A to A'. Substituting m for $\frac{Y_1' - Y_1}{X_1' - X_1}$ in (2.14) gives

$$m > -\cot\theta \tag{2.15}$$

which is the same as equation (2.9).

It is easy to see from Figure 6 that if this condition is violated, screen 1 and screen 2 will see opposite sides of the line joining A to A'. This motivates the following definition of the forbidden zone:

•*A point A' is in the forbidden zone of a point A if, and only if, the two screens see opposite sides of the straight line joining A' to A.*

As we have shown above, the forbidden zone is precisely the region where the ordering constraint breaks down.

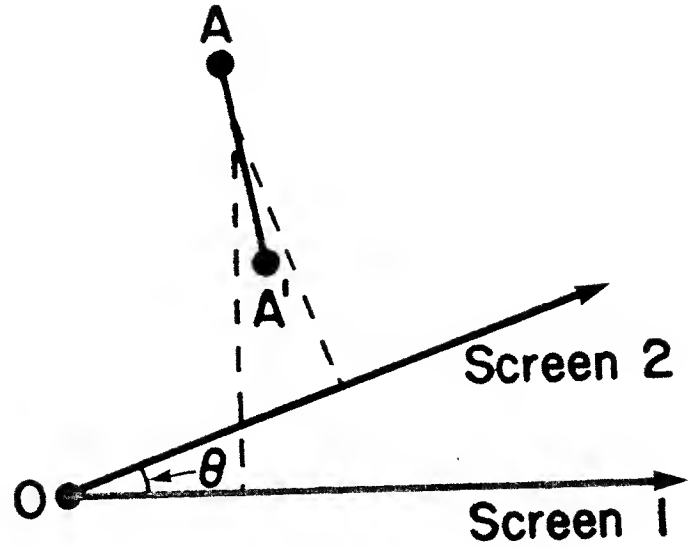

Figure 6 *See text.*

Observe from Figure 6 that the definition of the forbidden zone is symmetric, that is, if A' is in the forbidden zone of A, then A' is in the forbidden zone of A. Note also that the forbidden zone of A depends only on the position of A and the angle θ between the two screens. This will not be the case when we consider perspective projection. In Section 3, we prove some results about the forbidden zone for perspective projection, which will also apply to orthographic projection.

3. Perspective Projection

We now consider the ordering constraints for perspective projection.

The geometry is summarized in Figure 7. The two screens have unit normals $\underline{\alpha}_1$ and $\underline{\alpha}_2$ and focal length m. The foci have position vectors $\mathbf{f}_1$, and $\mathbf{f}_2$ relative to the origin 0. An arbitrary point A in space has position vector $\mathbf{X}$ relative to 0.

The centers of screen 1 and screen 2 are $\mathbf{r}_1^0$ and $\mathbf{r}_2^0$, where

$$\begin{aligned} \mathbf{r}_1^0 &= \mathbf{f}_1 - m\underline{\alpha}_1 \\ \mathbf{r}_2^0 &= \mathbf{f}_2 - m\underline{\alpha}_2 \end{aligned} \tag{3.1}$$

The equations of the screens are

$$\begin{aligned} \mathbf{r} \cdot \underline{\alpha}_1 &= \mathbf{f}_1 \cdot \underline{\alpha}_1 - m, \quad \text{for screen 1} \\ \mathbf{r} \cdot \underline{\alpha}_2 &= \mathbf{f}_2 \cdot \underline{\alpha}_2 - m, \quad \text{for screen 2.} \end{aligned} \tag{3.2}$$

Now consider the projection of a point $\mathbf{X}$ on screen 1. The line of projection is

$$\mathbf{r}(\lambda) = \mathbf{X} + \lambda(\mathbf{f}_1 - \mathbf{X}) \tag{3.3}$$

From (3.2), this hits screen 1 at

$$\mathbf{X} \cdot \underline{\alpha}_1 + \lambda(\mathbf{f}_1 - \mathbf{X}) \cdot \underline{\alpha}_1 = \mathbf{f}_1 \cdot \underline{\alpha}_1 - m \tag{3.4}$$

We solve (3.4) to obtain

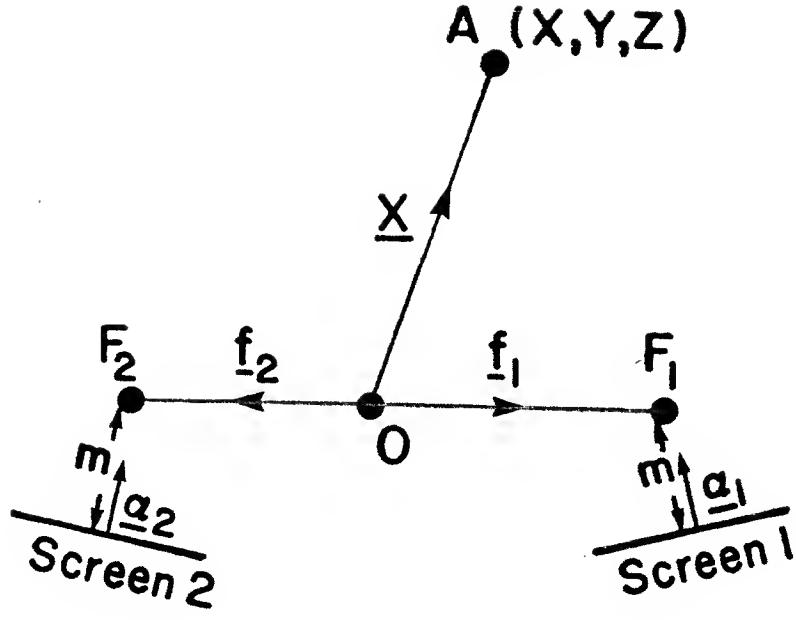

Figure 7 *See text.*

$$\lambda - 1 = \frac{-m}{(\mathbf{f}_1 - \mathbf{X}) \cdot \underline{\alpha}_1} \tag{3.5}$$

and so $\mathbf{X}$ is projected to a point $\mathbf{r}_1$ with

$$\mathbf{r}_1 = \mathbf{f}_1 - \frac{m(\mathbf{f}_1 - \mathbf{X})}{(\mathbf{f}_1 - \mathbf{X}) \cdot \underline{\alpha}_1} \tag{3.6}$$

To find the position $\mathbf{X}_1$ of the image of $\mathbf{X}$ relative to the center of the screen, rather than relative to 0, we must subtract $\mathbf{r}_1^0$ from $\mathbf{r}_1$. Hence,

$$\mathbf{X}_1 = \mathbf{r}_1 - \mathbf{r}_1^0 = \frac{m}{(\mathbf{f}_1 - \mathbf{X}) \cdot \underline{\alpha}_1} \langle \langle (\mathbf{f}_1 - \mathbf{X}) \cdot \underline{\alpha}_1 \rangle \underline{\alpha}_1 - (\mathbf{f}_1 - \mathbf{X}) \rangle \tag{3.7}$$

Similarly, the projection on screen 2 is given by

$$\mathbf{X}_2 = \frac{m}{(\mathbf{f}_2 - \mathbf{X}) \cdot \underline{\alpha}_2} \langle \langle (\mathbf{f}_2 - \mathbf{X}) \cdot \underline{\alpha}_2 \rangle \underline{\alpha}_2 - (\mathbf{f}_2 - \mathbf{X}) \rangle \tag{3.8}$$

Now we restrict ourselves to the geometry used by Longuet-Higgins (1982). In this case, the vectors $\underline{\alpha}_1, \underline{\alpha}_2, \mathbf{f}_1$ and $\mathbf{f}_2$ are coplanar and perpendicular to a vector k that we take to be in the z-direction. We can choose the origin 0 and the x-axis so that

$$\mathbf{f}_1 = -\mathbf{f}_2 = (f, 0, 0) \tag{3.9}$$

We define angles ϕ_1 and ϕ_2, such that

$$\begin{aligned} \underline{\alpha}_1 &= (\cos\phi_1, \sin\phi_1, 0) \\ \underline{\alpha}_2 &= (\cos\phi_2, \sin\phi_2, 0), \end{aligned} \tag{3.10}$$

and two vectors $\underline{\alpha}_1^*$ and $\underline{\alpha}_2^*$ orthogonal to $\underline{\alpha}_1$ and $\underline{\alpha}_2$

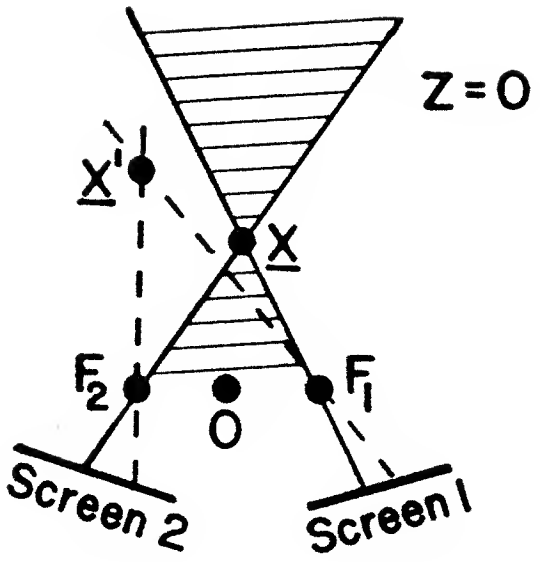

Figure 8 *See text.*

$$\begin{aligned} \underline{\alpha}_1^* &= (\sin\phi_1, -\cos\phi_1, 0) \\ \underline{\alpha}_2^* &= (\sin\phi_2, -\cos\phi_2, 0) \end{aligned} \tag{3.11}$$

The vectors $\underline{\alpha}_1^*$, $\underline{\alpha}_1$ and $\mathbf{k}$ form a right-handed triad for screen 1, as do $\underline{\alpha}_2^*$, $\underline{\alpha}_2$ and $\mathbf{k}$ for screen 2.

Let the components of $\mathbf{X}$ be (X, Y, Z). Its projection $\mathbf{X}_1$ on screen 1 is given by (3.7), and can be written

$$\mathbf{X}_1 = (\mathbf{X}_1 \cdot \underline{\alpha}_1^*)\alpha_1^* + (\mathbf{X}_1 \cdot \mathbf{k})\mathbf{k} \tag{3.12}$$

Note $\mathbf{X}_1 \cdot \underline{\alpha}_1 = 0$, since $\mathbf{X}_1$ lies on the screen. Thus we can take $X_1 = (\mathbf{X}_1 \cdot \underline{\alpha}_1^*)$ and $Z_1 = (\mathbf{X}_1 \cdot \mathbf{k})$ to be the cartesian coordinates of the image on screen 1. Substituting from (3.11), (3.9) and (3.7) gives

$$\begin{aligned} X_1 &= \frac{m}{f\cos\phi_1 - X\cos\phi_1 - Y\sin\phi_1}(-f\sin\phi_1 + X\sin\phi_1 - Y\cos\phi_1) \\ Z_1 &= \frac{mZ}{f\cos\phi_1 - X\cos\phi_1 - Y\sin\phi_1} \end{aligned} \tag{3.13}$$

Similarly, we have

$$\begin{aligned} X_2 &= \frac{-m}{f\cos\phi_2 + X\cos\phi_2 + Y\sin\phi_2}(f\sin\phi_2 + X\sin\phi_2 - Y\cos\phi_2) \\ Z_2 &= \frac{-mZ}{f\cos\phi_2 + X\cos\phi_2 + Y\sin\phi_2} \end{aligned} \tag{3.14}$$

It is well-known that there is an ordering constraint along epipolar lines, and with our geometry, one such line occurs when $Z = 0$. This is illustrated by Figure 8, where $\mathbf{X}$ and $\mathbf{X}'$ lie in the $Z = 0$ plane. The forbidden zone of $\mathbf{X}$ is the shaded region. Then, provided that $\mathbf{X}'$ lies outside the forbidden zone of $\mathbf{X}$ the images of $\mathbf{X}'$ on the two screens will be either both to the right of the image of $\mathbf{X}$, as in Figure 8, or both to the left.

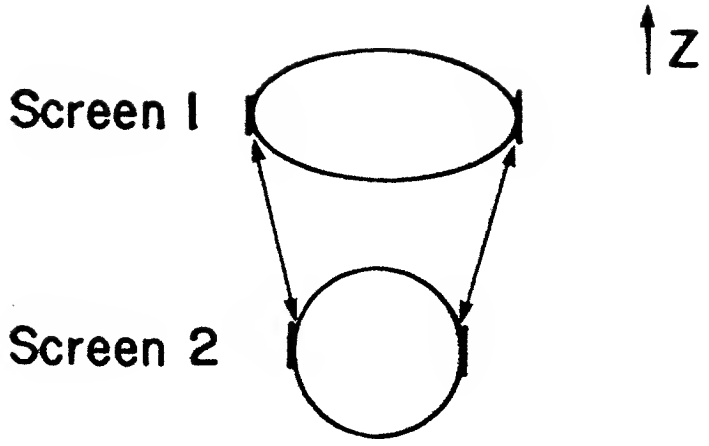

Figure 9 *See text.*

Now observe that both X_1 and X_2 in (3.13) and (3.14) are independent of Z. Thus, if we remove the restriction $Z = 0$, we obtain a similar ordering constraint *independent of whether* X *and* X′ *lie on the same epipolar line.* Intuitively, if X is to the "right" of X′ in space, then the projections of X will be to the "right" of the projections of X′ on both screens.

These results hold if, and only if, X′ is not in the forbidden zone of X. The forbidden zone is complicated for perspective projections, and we discuss it in the next subsection.

One obvious application of our generalized ordering constraint concerns two contours on the two screens. It implies that there is an ordered mapping between the two contours and that the points at the end of the contours, where $\frac{dX}{dZ} = 0$, must be identified. (See Figure 9.) The ordering constraint along epipolar lines arises as a special case when we take the contours to be the epipolar lines.

The Forbidden Zone for Perspective Projection.

Since the X-component on the screen is independent of the Z-component in space, (3.13) and (3.14), the shape of the forbidden zone will be independent of Z.

From Figure 10, it is clear that B' is in the forbidden zone of A if and only if the two screens see opposite sides of the straight line joining A to B'.

The structure of the forbidden zone is more complicated for perspective than for orthographic projection. For orthographic projection, the angle subtended by A to the two screens was a fixed angle θ, independent of the position of A. For perspective, this is no longer the case. In this section, we characterize the forbidden zone algebraically and prove various desirable properties about it.

Let the permissive zone of A be the complement of the forbidden zone of A. If B is in the permissive zone of A to the left of A we will write B *left* $*A$. Similarly, B *right* $*A$ means B is in the permissive zone of A to the right of A.

It is clear from Figure 10 that B *left* $*A$ if and only if the angles ψ_1 and ψ_2 are both positive. This will be true if and only if the Z-component of the cross-product of the vectors $\underline{F_2A}$ with $\underline{F_2B}$ and $\underline{F_1A}$ with $\underline{F_1B}$ are both positive (since the lines are coplanar, the sign of the Z-component is the sign of $\sin\psi_2$ and $\sin\psi_1$, respectively).

The vectors $\underline{F_2A}$ and $\underline{F_2B}$ are written

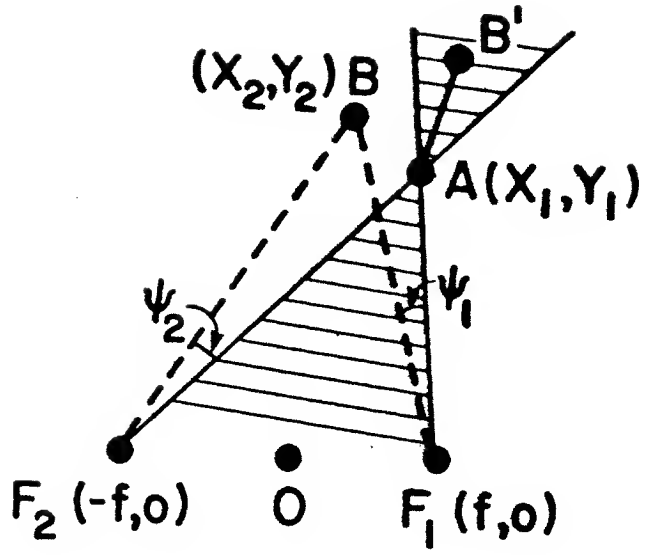

Figure 10 *B' lies in the forbidden zone of A but B does not.*

$$\begin{aligned} \underline{F_2A} &= (X_1 + f, Y_1, 0) \\ \underline{F_2B} &= (X_2 + f, Y_2, 0) \end{aligned} \tag{3.15}$$

and their cross-product is

$$\underline{F_2A} \times \underline{F_2B} = \{(X_1 + f)Y_2 - (X_2 + f)Y_1\}\mathbf{k} \tag{3.16}.$$

Thus the point B is to the left of the line F_2A if

$$(X_1 + f)Y_2 - (X_2 + f)Y_1 > 0 \tag{3.17}$$

and to the right if

$$(X_1 + f)Y_2 - (X_2 + f)Y_1 < 0. \tag{3.18}$$

Similarly, B is to the left of F_1A if

$$(X_1 - f)Y_2 - (X_2 - f)Y_1 > 0 \tag{3.19}$$

and to the right if

$$(X_1 - f)Y_2 - (X_2 - f)Y_1 < 0. \tag{3.20}$$

So we can have B *left* $*A$ if and only if both

$$\begin{aligned} (X_1 + f)Y_2 - (X_2 + f)Y_1 &> 0 \\ (X_1 - f)Y_2 - (X_2 - f)Y_1 &> 0 \end{aligned} \tag{3.21}$$

and B *right*$*A$ if and only if both

$$\begin{aligned} (X_1 + f)Y_2 - (X_2 + f)Y_1 &< 0 \\ (X_1 - f)Y_2 - (X_2 - f)Y_1 &< 0 \end{aligned} \tag{3.22}$$

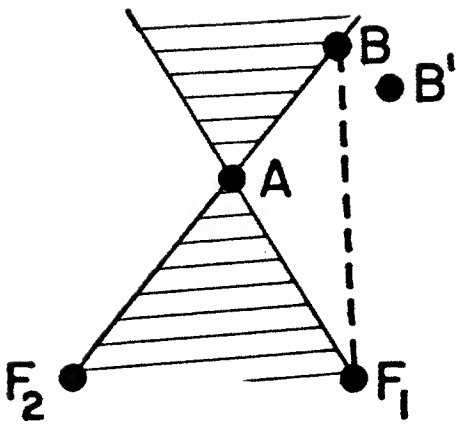

Figure 11 *The line AB projects to a line on screen 1 and a point on screen 2.*

Observe that if $Y_2 > Y_1$ then if B is to the left of F_1A it is automatically to the left of F_2A. If $Y_2 < Y_1$, then if B is to the left of F_2A it is automatically to the left of F_1A. This is an alternative way of characterizing *left** and *right**.

We now show that the forbidden zone has the desirable properties of symmetry and transitivity.

For symmetry, we must show that B *left**A is equivalent to A *right**B. In other words, if B is in the permissive zone of A to the left, then A is in the permissive zone of B to the right and conversely.

The conditions for B *left**A are given by (3.21). The condition for A *right**B are, from (3.22)

$$\begin{aligned}(X_2 + f)Y_1 - (X_1 + f)Y_2 < 0 \\ (X_2 - f)Y_1 - (X_1 - f)Y_2 < 0\end{aligned} \tag{3.23}$$

The equation (3.21) and (3.23) are clearly equivalent and hence the result is proved.

Now we consider transitivity. We want to show that if C left*B and B left*A then C left*A.

C left*B means that F_2C is to the left of F_2B and F_1C is to the left of F_1B. From B left*A we have F_2B to the left of F_2A and F_1B is to the left of F_1A. So, F_2C is to the left of F_2B which is to the left of F_2A and hence, F_2C is to the left of F_2A. Similarly, F_1C is to the left of F_1A. Thus C left*A and the result holds.

4. Limits to Fusion

We have shown in the preceding sections that when a point B is inside the forbidden zone of a point A, different sides of the straight line joining A to B are projected to different screens. If B lies on the boundary of the forbidden zone of A then, as in Figure 11, one side of the line will be visible to one screen but the other screen will only see a point. This corresponds to Panum's Limiting case.

If the point B is moved to B' in the permissive region of A then the line AB' can be seen by both screen 1 and screen 2 although the projection to screen 2 will be very foreshortened. It should in principle be possible to fuse features on the line AB' despite this foreshortening.

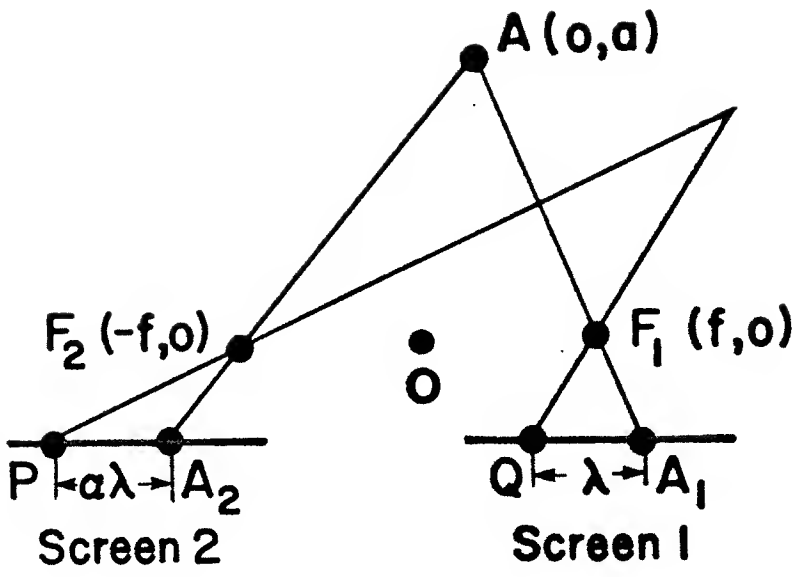

Figure 12 *See text*

In a recent paper, however, Burt and Julesz (1980) claim that there is a limit to the relative disparities in two screens that the human visual system can fuse. If this is so, then the boundaries of the forbidden zone may be a competence limit while the Burt-Julesz limit may be a performance limit. We have so far been assuming that the greatest slopes in space that we can see are bounded above by the forbidden zone and Burt-Julesz' results could tell us that this limit cannot always be attained. To investigate this, we ask what surfaces in space correspond to a limit of disparity gradient. We will see that they are straight lines.

First, assume the system is fixating at infinity so the two screens are parallel. We illustrate this case by figure (12). Now let A be a point in the center of the visual field with coordinate $(0, a)$. It projects to points A_1 and A_2 on the screens. These have components $(f(1+\frac{m}{a}), -m)$ and $(-f(1+\frac{m}{a}), -m)$ respectively. Let the relative disparity limit be denoted by α (Burt-Julesz would set $\alpha = 3$). The relative disparity of two points in space is defined as the ratio of the differences of their images in the two screens (see Fig. 12). So the fusional limit would occur for points Q and P on the two screens, where Q is a distance λ to the left of A_1 and P is $\alpha\lambda$ to the left of A_2. So we have

$$\begin{aligned} \underline{OQ} &= (f(1+\frac{m}{a}) - \lambda, -m) \\ \underline{OP} &= (-f(1+\frac{m}{a}) - \alpha\lambda, -m) \end{aligned} \tag{4.1}$$

The line QF_1 is given by

$$y = \frac{m}{\lambda - \mu}(x - f) \tag{4.2}$$

where $\mu = mf/a$, and the line PF_2 is

$$y = \frac{m}{\alpha\lambda + \mu}(x + f). \tag{4.3}$$

These lines intersect at a point where

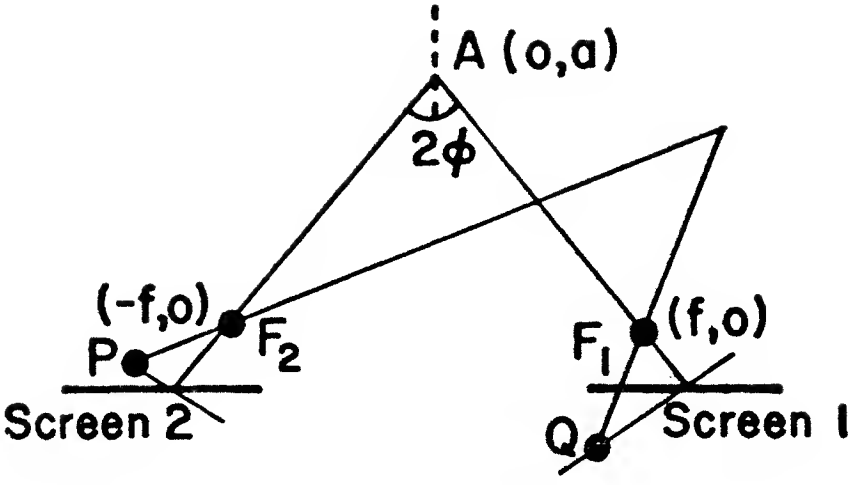

Figure 13 *The screens fixate at A with vergence angle 2φ*

$$\begin{aligned} x &= \frac{f\lambda(1+\alpha)}{2\mu+\lambda(\alpha-1)} \\ y &= \frac{2mf}{2\mu+\lambda(\alpha-1)} \end{aligned} \tag{4.4}$$

Letting λ vary, we obtain a curve passing through A which is the greatest slope corresponding to the Burt-Julesz limit. Eliminating λ from (4.4) gives the curve

$$ax = \left(\frac{\alpha+1}{\alpha-1}\right) f(a-y) \tag{4.5}$$

which is a straight line. Observe that the gradient of the curve depends only on a and f. Setting $\alpha = 3$ gives a gradient of $\frac{-a}{2f}$.

We also perform the calculation when the visual system is fixating on A as in Figure 13.

It is straightforward to modify the previous calculation by setting

$$\begin{aligned} m' &= m\cos\phi \\ \lambda' &= \lambda\cos\phi \end{aligned} \tag{4.6}$$

Note that in this case we have

$$\tan\phi = \frac{f}{a}. \tag{4.7}$$

Since neither m or λ appear in the final expression (4.5), the result is unaltered by the change in fixation point. As before, the gradient is proportional to $\frac{a}{f}$ and by (4.7) it is hence proportional to $\cos\phi$.

5. The generalized ordering constraint and other constraints

Arnold and Binford (1980) and Mayhew and Frisby (1981) suggested a figural continuity constraint for stereo matching: disparity is usually continuous along contours. This figural continuity constraint can be powerful in practical implementations as demonstrated by

Grimson (1984) and Ohta and Kanade (1984). Our geometric model of matching implicitely implements figural continuity, because matching is performed along continuous contours under the constraint that the disparity gradient is not too large (i.e., the contour does not enter the forbidden zone).

Binford (see Baker et al., 1983) suggested surface occlusion rules for making explicit opacity and non-opacity of surfaces. In particular the cross-product rule determines whether an hypothetized match between two corners lies in the forbidden zone. Again a rule such as this follows immediately from the generalized ordering constraint.

6. Epipolar lines as a constraint for matching

We now proceed to show that if the stereo geometry is known, perhaps by registering the screens from the image as in Appendix 2, then there is a simple relation between the epipolar lines on the two screens. A point on the left screen will lie on a unique epipolar line which will therefore correspond to a unique epipolar line in the right screen.

We consider a general point in space (X, Y, Z). On the right screen, this projects to (X_1, Z_1) by equation 3.13,

$$X_1 = \frac{m}{f\cos\phi_1 - X\cos\phi_1 - Y\sin\phi_1}(-f\sin\phi_1 + X\sin\phi_1 - Y\cos\phi_1) \tag{6.1}$$

$$Z_1 = \frac{mZ}{f\cos\phi_1 - X\cos\phi_1 - Y\sin\phi_1} \tag{6.2}$$

and on the left screen to (X_2, Z_2) by (3.14),

$$X_2 = \frac{-m}{f\cos\phi_2 + X\cos\phi_2 + Y\sin\phi_2}(f\sin\phi_2 + X\sin\phi_2 - Y\cos\phi_2) \tag{6.3}$$

$$Z_2 = \frac{-mZ}{f\cos\phi_2 + X\cos\phi_2 + Y\sin\phi_2} \tag{6.4}$$

To derive our relation, we use the four equations (6.1)-(6.4) to eliminate the unknowns X, Y and Z, which leaves us with a single equation relating X_1, X_2, Z_1 and Z_2.

We eliminate Z by dividing (6.2) by (6.4) to obtain

$$\frac{Z_1}{Z_2} = \frac{-\{f\cos\phi_2 + X\cos\phi_2 + Y\sin\phi_2\}}{f\cos\phi_1 - X\cos\phi_1 - Y\sin\phi_1} \tag{6.5}$$

We rewrite (6.1) and (6.3) respectively as

$$\begin{aligned} &X\{-X_1\cos\phi_1 - m\sin\phi_1\} \\ &\quad + Y\{-X_1\sin\phi_1 + m\cos\phi_1\} \\ &\quad + X_1 f\cos\phi_1 + mf\sin\phi_1 = 0 \end{aligned} \tag{6.6}$$

and

$$\begin{aligned} &X\{-X_2\cos\phi_2 - m\sin\phi_2\} \\ &\quad + Y\{-X_2\sin\phi_2 + m\cos\phi_2\} \\ &\quad - X_2 f\cos\phi_2 - mf\sin\phi_2 = 0 \end{aligned} \tag{6.7}$$

We combine (6.6) and (6.7) into a matrix equation

$$\begin{pmatrix} -X_1\cos\phi_1 - m\sin\phi_1 & -X_1\sin\phi_1 + m\cos\phi_1 \\ -X_2\cos\phi_2 - m\sin\phi_2 & -X_2\sin\phi_2 + m\cos\phi_2 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} -X_1 f\cos\phi_1 & -mf\sin\phi_1 \\ X_2 f\cos\phi_2 & +mf\sin\phi_2 \end{pmatrix} \quad (6.8)$$

We invert the matrix to solve for X and Y obtaining

$$X = \frac{f\{(X_1X_2 - m^2)\sin(\phi_1+\phi_2) - m(X_1+X_2)\cos(\phi_1+\phi_2)\}}{\{X_1X_2+m^2\}\sin(\phi_2-\phi_1) + m(X_2-X_1)\cos(\phi_2-\phi_1)} \quad (6.9)$$

and

$$Y = -2f\frac{\{X_1X_2\cos\phi_2\cos\phi_1 + mX_2\cos\phi_2\sin\phi_1 + mX_1\sin\phi_2\cos\phi_1 + m^2\sin\phi_2\sin\phi_1\}}{\{X_1X_2+m^2\}\sin(\phi_2-\phi_1) + m(X_2-X_1)\cos(\phi_2-\phi_1)} \quad (6.10)$$

We rewrite (6.5) as

$$\begin{aligned}(Z_1\cos\phi_1 + Z_2\cos\phi_2)f + (Z_2\cos\phi_2 - Z_1\cos\phi_1)X \\ + (Z_2\sin\phi_2 - Z_1\sin\phi_1)Y = 0.\end{aligned} \quad (6.11)$$

and substitute (6.9) and (6.10) into (6.11). After some manipulation, this simplifies and we obtain the result

$$Z_1\{m\sin\phi_2 + X_2\cos\phi_2\} = Z_2\{m\sin\phi_1 + X_1\cos\phi_1\} \quad (6.12)$$

Thus we see that if we know the position on one screen, for example (X_1, Z_1), we find that the other position on the other screen, (X_2, Z_2), must lie on a straight line. This is, of course, the standard epipolar line constraint. It is interesting to see that in our coordinate system, it can be written in such a simple form. This makes it straightforward to implement it in an algorithm.

7. Outline of a contour based algorithm

A simple example of an algorithm which indicates how the generalized ordering constraint can be exploited is as follows.

Instead of using the epipolar lines to impose an ordering constraint, we use them as a consistency constraint. In conjunction with the ordering constraint along the contour, they determine stereo matching. This avoids wasteful scanning of all epipolar lines and ensures that computation is done only at places where it is necessary. It also automatically enforces figural continuity. There are clearly interesting effects when the contour we are matching runs along an epipolar line for some distance. Psychophysical experiments (Buelthoff and Poggio, pers. comm.) suggest that the human visual system also has problems in such cases.

For a single contour the generalized ordering constraint tells us that the leftmost point and the rightmost point in the two images must correspond (see figure 9). These points can be used to "register" the viewing system, using the formulae in Appendix 2, and hence to determine implicitely the epipolar lines. Starting at one of these two points we move along the contours using the epipolar line constraint (6.12) to determine which points are to be matched. If more than one contour is present the generalized ordering constraint is used to decide which contours in the two images correspond (contours are also "ordered" in a similar way as points along a single contour are). In this case there will be many points in the two images which are known to correspond (two for each contour) and the registration process will hence be more robust (Mayhew and Longuet-Higgins, 1982). An algorithm of this type is currently under development with E. Tiffany.

8. Discussion

A few remarks about photometric effects and perspective invariants are relevant at this stage. We have assumed that the contours in the image being matched correspond to a physical structure, such as a wire or a silhouette, in the object being viewed. If zero-crossings of the Laplacian of a Gaussian are used as the matching primitives we must be sure they correspond to a precise physical location in the object being viewed. It is encouraging that results by Yuille (1984) suggest that many zero-crossings may be due to significant changes in the geometry of the object which are almost independent of the viewing positions. As it stands now, the theory presented here is only valid for wire-frame objects. Occluding contours, though theoretically "wrong", may however be often used in small angle stereo (Grimson, 1981). The main challenge is to extend the matching scheme presented here to solid and textured surfaces. Several possibilities can be considered: a) fingerprints representations (Yuille and Poggio, 1983) may provide specific features to be used in the matching process; b) several functional measurements of the two images (Krass, 1984) along the contour may be used to perform the matching.

The generalized ordering constraint analysis is done under the assumption of a three-dimensional curve in space. The analysis may be applied to occluding contours, but we need estimates of the errors. We also need to identify which contours are occluding contours. One possibility is to use fingerprints (Yuille and Poggio, 1983). The idea would then be to identify separately zero crossings corresponding to occluding contours, zero crossings corresponding to step edges and zero crossings due to texture. Some of these zero-crossings could then be used for a matching scheme based on the generalized ordering constraint. Interpolation of the surface between matched zero crossing could be performed under the constraint of matching measurements provided by a scheme similar to Kass' (1984). His scheme may be reformulated to exploit a form of ordering constraint for speeding up the search. The main problem for a practical algorithm is obviously the stability of the contours to be matched between the two images. Geometric and photometric distortions are likely to present a hard problem.

It is also possible to use perspective invariants to help match two contours which arise from the same physical location. Points on the two contours which correspond to the same invariant can be matched. A number of perspective invariants or "semi-invariants" are discussed by Yuille and Verri (1984). For example, the zeros and the discontinuities of curvature of a non-planar curve are preserved under perspective projection to the image plane, although the converse is not true. These "semi-invariants" can be used to find points on two contours which correspond.

In summary, we have shown that, with the Longuet-Higgins (1982) geometry, a generalized ordering constraint holds provided the viewed object does not enter the *forbidden zone*. This constraint is not restricted to epipolar lines. We characterize the forbidden zone for orthographic and perspective projections and discuss experiments which suggest that the human visual system incorrectly interprets objects in this zone. Results by Burt and Julesz (1980) suggest that limits to fusion occur before this zone is entered. We use the generalized ordering constraint to propose an algorithm for stereo matching along contours using the epipolar lines as a consistency constraint. We use the generalized ordering constraint to determine points in the images which correspond and hence can be used to find the viewing parameters. The generalized ordering constraint provides an efficient way to match points by scanning only along the available contours instead of along all epipolar lines. It also automatically imposes a figural continuity constraint (Mayhew, 1983).

9. Acknowledgements

Some of the motivation for this work was provided by Vincent Torre. Mike Brady and Eric Grimson made useful comments on the manuscript. Carol Bonomo would like to be acknowledged here.

Appendix 1: The Orthographic Limit of Perspective Projection

From equation (3.7), we have the perspective projection of a point X on a screen with normal $\underline{\alpha}_1$, focus $\mathbf{f}_1$ and focal length m is given by $\mathbf{x}_1$ where:

$$\mathbf{X}_1 = \frac{-m}{(\mathbf{f}_1 - \mathbf{X}) \cdot \underline{\alpha}_1}(\{(\mathbf{f}_1 - \mathbf{X}) - \{(\mathbf{f}_1 - \mathbf{X}) \cdot \underline{\alpha}_1\}\underline{\alpha}_1\}) \tag{A.1.1}$$

The term inside the brackets corresponds to the orthographic projection of $(\mathbf{f}_1 - \mathbf{X})$ on screen 1. It is scaled by a term $(\mathbf{f}_1 - \mathbf{X}) \cdot \underline{\alpha}_1$. Provided we stay in a region of the image where this term only varies a little, orthographic projection will be a good approximation and the scaling will be constant. This will normally be the case when $(\mathbf{f} - \mathbf{X})$ and $\underline{\alpha}_1$ are almost parallel, since we have

$$(\mathbf{f} - \mathbf{X}) \cdot \underline{\alpha}_1 = |\mathbf{f} - \mathbf{X}| \cos\tau \tag{A.1.2}$$

where τ is the angle between $(\mathbf{f} - \mathbf{X})$ and $\underline{\alpha}_1$. For τ near zero, $\cos\tau$ is approximated by $1 - \tau^2/2$ and will be insensitive to small changes in τ.

How good the orthographic approximation is in general will depend on the gradient of the surface being projected not being too large.

Appendix 2: Solving for the Viewing Angles

The projection $\mathbf{X}_1$ of a point X on screen 1 is given by equation (3.7), and we write it as

$$\mathbf{X}_1 = m\underline{\alpha}_1 - \frac{m}{\mathbf{X} - \mathbf{f}_1 \cdot \underline{\alpha}}(\mathbf{X} - \mathbf{f}_1) \tag{A.2.1}$$

Alternately, we can take a point P_1 on screen 1 and construct the line joining it to the focus (F_1). The point P_1 has position vector $\mathbf{Z}_1$ relative to the origin $\mathbf{X}_o$ of screen 1. From Figure (1) and Figure (8)we see that

$$\mathbf{X}_o = \mathbf{f}_1 - m\underline{\alpha}_1 \tag{A.2.2}$$

The gradient of the line from P_1 to F_1 is $m\underline{\alpha}_1 - \mathbf{Z}_1$, and we can write the line as

$$\mathbf{r}(\lambda) = \mathbf{f}_1 - m\underline{\alpha}_1 + \mathbf{Z}_1 + \lambda\{m\underline{\alpha}_1 - \mathbf{Z}_1\} \tag{A.2.3}$$

We rewrite (A.2.3) as

$$\mathbf{r}(\lambda) = \mathbf{f}_1 + (\lambda - 1)\{m\underline{\alpha}_1 - \mathbf{Z}_1\} \tag{A.2.4}$$

and we obtain a similar equation for screen 2:

$$\mathbf{r}(\mu) = \mathbf{f}_2 + (\mu - 1)\{m\underline{\alpha}_2 - \mathbf{Z}_2\} \tag{A.2.5}$$

These lines intersect at a point P where

$$\mathbf{f}_1 + (\lambda - 1)\{m\underline{\alpha}_1 - \mathbf{Z}_1\} = \mathbf{f}_2 + (\mu - 1)\{m\underline{\alpha}_2 - \mathbf{Z}_2\} \qquad (A.2.6)$$

Now in our geometry there is a fixed (vertical) axis k such that

$$\underline{\alpha}_1 \cdot \mathbf{k} = \underline{\alpha}_2 \cdot \mathbf{k} = \mathbf{f}_1 \cdot \mathbf{k} = \mathbf{f}_2 \cdot \mathbf{k} = 0 \qquad (A.2.7)$$

P_1 lies on screen 1 so $\mathbf{Z}_1 \cdot \underline{\alpha}_1 = 0$ and so we can define coordinates $Z_1 (= \mathbf{Z}_1 \cdot \mathbf{k})$ and X_1 of $\mathbf{Z}_1$ by

$$\mathbf{Z}_1 = Z_1\mathbf{k} + X_1(\mathbf{k} \times \underline{\alpha}_1) \qquad (A.2.8)$$

Similarly, for $\mathbf{Z}_2$ we have

$$\mathbf{Z}_2 = Z_2\mathbf{k} + X_2(\mathbf{k} \times \underline{\alpha}_2) \qquad (A.2.9)$$

Taking the dot product of (A.2.6) with k yields

$$0 = -(\mu - 1)Z_2 + (\lambda - 1)Z_1 \qquad (A.2.10)$$

For convenience we replace λ and μ with ς and ν where

$$\begin{aligned} \varsigma &= \lambda - 1 \\ \nu &= \mu - 1 \end{aligned} \qquad (A.2.11)$$

Then we rewrite (A.2.6) using (A.2.10) as

$$\begin{aligned} \mathbf{f}_1 - \mathbf{f}_2 = \nu(m\underline{\alpha}_2 - X_2(\mathbf{k} \times \underline{\alpha}_2)) \\ - \varsigma(m\underline{\alpha}_1 - X_1(\mathbf{k} \times \underline{\alpha}_1)) \end{aligned} \qquad (A.2.12)$$

and equation (A.2.10) as

$$\frac{\varsigma}{\nu} = \frac{Z_2}{Z_1} \qquad (A.2.13)$$

Substituting for ς from (A.2.13) into (A.2.12) gives

$$\mathbf{f}_1 - \mathbf{f}_2 = \nu\{(m\underline{\alpha}_2 - X_2(\mathbf{k} \times \underline{\alpha}_2)) - \frac{Z_2}{Z_1}(m\underline{\alpha}_1 - X_1(\mathbf{k} \times \underline{\alpha}_1))\} \qquad (A.2.14)$$

Now we assume a second point is also known to correspond. This has coefficients $\overline{Z}_1, \overline{X}_1$ on screen 1 and $\overline{Z}_2, \overline{X}_2$ on screen 2. Repeating the argument leading up to (A.2.14) gives the equation

$$\mathbf{f}_1 - \mathbf{f}_2 = \tau\{(m\underline{\alpha}_2 - \overline{X}_2(\mathbf{k} \times \underline{\alpha}_2)) - \frac{\overline{Z}_2}{\overline{Z}_1}(m\underline{\alpha}_1 - \overline{X}_1(\mathbf{k} \times \underline{\alpha}_1))\} \qquad (A.2.15)$$

for some constant τ.

Define $2\mathbf{f} = \mathbf{f}_1 - \mathbf{f}_2$ and the angles θ and ϕ as in figure (14). Then taking the dot product of (A.2.15) with $\underline{\alpha}_1$ and $\underline{\alpha}_2$ yields

Figure 14 *See text*

$$\frac{1}{\nu}2f\cos(\theta+\phi) = m\cos\theta - X_2\sin\theta - \frac{Z_2}{Z_1}m \tag{A.2.16}$$

and

$$\frac{1}{\nu}2f\cos\phi = m - \frac{Z_2}{Z_1}(m\cos\theta + X_1\sin\theta) \tag{A.2.17}$$

Similarly from (A.2.15) we can obtain

$$\frac{1}{\tau}2f\cos(\theta+\phi) = m\cos\theta - \overline{X}_2\sin\theta - \frac{\overline{Z}_2}{\overline{Z}_1}m \tag{A.2.18}$$

and

$$\frac{1}{\tau}2f\cos\phi = m - \frac{\overline{Z}_2}{\overline{Z}_1}(m\cos\theta + \overline{X}_1\sin\theta) \tag{A.2.19}$$

Dividing (A.2.16) by (A.2.18) yields

$$\frac{\tau}{\nu} = \frac{m\cos\theta - X_2\sin\theta - \frac{Z_2}{Z_1}m}{m\cos\theta - \overline{X}_2\sin\theta - \frac{\overline{Z}_2}{\overline{Z}_1}m} \tag{A.2.20}$$

Similarly, dividing (A.2.17) by (A.2.19) yields

$$\frac{\tau}{\nu} = \frac{m - \frac{Z_2}{Z_1}(m\cos\theta + X_1\sin\theta)}{m - \frac{\overline{Z}_2}{\overline{Z}_1}(m\cos\theta + \overline{X}_1\sin\theta)} \tag{A.2.21}$$

Hence we can combine (A.2.20) with (A.2.21) to yield an equation in θ only.

$$\begin{aligned}(m\cos\theta - X_2\sin\theta - \frac{Z_2}{Z_1}m)(m - \frac{\overline{Z}_2}{\overline{Z}_1}(m\cos\theta + \overline{X}_1\sin\theta)) = \\ (m\cos\theta - \overline{X}_2\sin\theta - \frac{\overline{Z}_2}{\overline{Z}_1}m)(m - \frac{Z_2}{Z_1}(m\cos\theta + X_1\sin\theta))\end{aligned} \tag{A.2.22}$$

Set

$$A = \frac{-Z_2}{Z_1}, \quad B = \frac{-\overline{Z}_2}{\overline{Z}_1} \tag{A.2.23}$$

Then, after some manipulation we write (A.2.22) as

$$\begin{aligned} &\{m^2(A-B) + (AX_1\overline{X}_2 - B\overline{X}_1X_2)\}\sin^2\theta \\ &\quad + m\{(\overline{X}_2 - X_2) + AB(\overline{X}_1 - X_1)\}\sin\theta \\ &\quad + m\{B(\overline{X}_1 - X_2) + A(\overline{X}_2 - X_1)\}\sin\theta\cos\theta = 0 \end{aligned} \tag{A.2.24}$$

We can divide this equation by $\sin\theta$ ($\sin\theta = 0$ corresponds to the special case when the eyes are fixating at infinity) and obtain an equation of form

$$C_1\sin\theta + C_2 + C_3\cos\theta = 0 \tag{A.2.25}$$

where

$$\begin{aligned} C_1 &= m^2(A-B) + (AX_1\overline{X}_2 - B\overline{X}_1X_2) \\ C_2 &= m\{\overline{X}_2 - X_2 + AB(\overline{X}_1 - X_1)\} \\ C_3 &= m\{B(\overline{X}_1 - X_2) + A(\overline{X}_2 - X_1)\} \end{aligned} \tag{A.2.26}$$

(A.2.25) can be written as a quadratic equation

$$(C_1^2 + C_3^2)\cos^2\theta + 2C_2C_3\cos\theta + (C_2^2 - C_1^2) = 0 \tag{A.2.27}$$

and has solution

$$\cos\theta = \frac{-C_2C_3 \pm C_1\sqrt{C_1^2 + C_3^2 - C_2^2}}{C_1^2 + C_3^2} \tag{A.2.28}$$

Note that $(C_1, C_2, C_3) \mapsto (-C_1, -C_2, -C_3)$ as $(A, B) \mapsto (B, A)$ and $(X_1, X_2, \overline{X}_1, \overline{X}_2) \mapsto (\overline{X}_1, \overline{X}_2, X_1, X_2)$. So if we have two points in the visual field which are known to coincide then we know that points P_1 and $\overline{P}_1$ on screen 1 correspond to P_2 and $\overline{P}_2$ on screen 2. The coefficients $(X_1, Z_1), (\overline{X}_1, \overline{Z}_1)$, (X_2, Z_2) and $(\overline{X}_2, \overline{Z}_2)$ of $P_1, \overline{P}_1, P_2$ and $\overline{P}_2$ are known and hence from (A.2.23) and (A.2.26), C_1, C_2 and C_3 are known. (A.2.28) gives two possible solutions for θ in terms of C_1, C_2 and C_3. Let these be θ_1 and θ_2. Dividing (A.2.16) by (A.2.17) yields

$$\frac{\cos(\theta + \phi)}{\cos\phi} = \frac{m\cos\theta - X_2\sin\theta - Z_2/Z_1 m}{m - Z_2/Z_1(m\cos\theta + X_1\sin\theta)} \tag{A.2.29}$$

Expanding $\cos(\theta + \phi) = \cos\theta\cos\phi - \sin\theta\sin\phi$ we can rewrite (A.2.29) as

$$\overline{C}_1\sin\phi + \overline{C}_3\cos\phi = 0 \tag{A.2.30}$$

where

$$\begin{aligned} \overline{C}_1 &= m - Z_2/Z_1(m\cos\theta + X_1\sin\theta) \\ \overline{C}_3 &= -X_2 - Z_2/Z_1(m\sin\theta - X_1\cos\theta) \end{aligned} \tag{A.2.31}$$

There are two possible values of $(\overline{C}_1, \overline{C}_2, \overline{C}_3)$ depending on whether $\theta = \theta_1$ or $\theta = \theta_2$. (A.2.31) will give two possible solutions for ϕ, for each value of θ:

$$\cos\phi = \frac{\pm\overline{C}_1\sqrt{\overline{C}_1^2 + \overline{C}_3^2}}{\overline{C}_1^2 + \overline{C}_3^2} \qquad (A.2.32)$$

Hence we have four possible solutions for θ and ϕ which can be formed directly from (A.2.28) and (A.2.29). In the general case, only one of these solutions will be physically reasonable and the others can be discarded.

We have shown that if the images on both screens of points in space are known then we can solve the non-linear equations directly for the angles θ and ϕ that specify the stereo geometry. These solutions are given by (A.2.28) and (A.2.32). Two solutions are generated of which only one is physically reasonable in general.

10. References

Arnold, R.D. and Binford, T.O. "Geometric constraints in stereo vision," *Proc. SPIE, San Diego* 238, (1980), 281-292.

Baker, H. H. "Depth from edge and intensity based stereo," Stanford University Technical Report STAN-CS-82-930, September, 1982.

Baker, H. H. and Binford, T. O. "Depth from Edge and Intensity Based Stereo," *Seventh International Joint Conference on Artificial Intelligence*, August 1981, 631-636.

Baker, H. H., Binford, T. O., Malik, J., Meller, J.-F. "Progress in Stereo Mapping" *Proc. Image Understanding Workshop*, SAI, L. Bauman, ed., June, 1983.

Burt, P. and Julesz, B. "A disparity gradient limit for binocular fusion" *Science*, 208, 9, 615-617, 1980.

Grimson, W.E.L. *From Images to Surfaces: A computational study of the human early visual system* MIT Press, Cambridge, Ma., 1981.

Grimson, W. E. L. "Computational Experiments with a Feature Based Stereo Algorithm", A. I. Memo 762, Mass. Institute of Technology, 1984.

Kass, M.H. "Computating stereo correspondance", *M. S., Mass. Institute of Technology*, 1984.

Krol, J. D. and van der Grind, W. A."The double-nail illusion: experiments on binocular vision with nails, needles and pins", *Perception*, 9, 651-669, 1980.

Krol, J. D. and van de Grind, W. A. "Rehabilitation of a classical notion of Panum's fusional area" *Perception*. 11, 615-619, 1982.

Longuet-Higgins, H. C. "The role of the vertical dimension in stereoscopic vision," *Perception* 11 (1982) 377-386.

Marr, D. and Poggio, T. "A theory of human stereo vision," *Proc. Roy. Soc. Lond.*B 204 (1979), 301-328. (an earlier version appeared as MIT AI Lab Memo 451, 1977).

Mayhew, J. E. W."Stereopsis", *Physical and Biological Processing of Images*, O. J. Braddick and A. C. Sleigh, eds., Springer-Verlag, Berlin, 1982.

Mayhew, J.E.W. and Frisby, J.P. "Psychophysical and computational studies towards a theory of human stereopsis," *Artificial Intelligence* 17 (1981), 349-385.

Mayhew, J.E.W. and Longuet-Higgins, H.C. "A computational model of binocular depth perception," *Nature Lond.* 297 (1982) 376-379.

Ohta, Y. and Kanade, T. "Stereo by intra- and inter-scanline search using dynamic programming," Carnegie-Mellon University Technical Report CMU-CS-83-162, 1983.

Verri, A. "Methodi Matematici per la Visione Stereografica," Ph.D. thesis, University of Genoa, 1984.

Yuille, A. "Zero-crossings on lines of Curvature," MIT AI Lab Memo 718, 1984.

Yuille, A. and Poggio, T. "Fingerprint Theorems for Zero Crossings" A. I. Memo 730, Mass. Institute of Technology, 1984.

Yuille, A. and Verri, A. In preparation